

A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Inventorship..... Gupta et al.  
Applicant..... Microsoft Corporation  
Attorney's Docket No. .... MS1-302US  
Title: Interactive Playlist Generation Using Annotations

TRANSMITTAL LETTER AND CERTIFICATE OF MAILING

To: Commissioner of Patents and Trademarks,  
Washington, D.C. 20231

From: Allan T. Sponseller (Tel. 509-324-9256; Fax 509-323-8979)  
Lee & Hayes, PLLC  
421 W. Riverside Avenue, Suite 500  
Spokane, WA 99201

The following enumerated items accompany this transmittal letter and are being submitted for the matter identified in the above caption.

1. Specification—title page, plus 45 pages, including 33 claims and Abstract
2. Transmittal letter including Certificate of Express Mailing
3. 8 Sheets Formal Drawings (Figs. 1-11)
4. Return Post Card

Large Entity Status [x]

Small Entity Status [ ]

Date: Sept. 15, 1999

By: [Signature]  
Allan T. Sponseller  
Reg. No. 38,318

CERTIFICATE OF MAILING

I hereby certify that the items listed above as enclosed are being deposited with the U.S. Postal Service as either first class mail, or Express Mail if the blank for Express Mail No. is completed below, in an envelope addressed to The Commissioner of Patents and Trademarks, Washington, D.C. 20231, on the below-indicated date. Any Express Mail No. has also been marked on the listed items.

Express Mail No. (if applicable) EL427262329

Date: Sept. 15, 1999

By: [Signature]  
Helen M. Hare

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**INTERACTIVE PLAYLIST GENERATION USING  
ANNOTATIONS**

Inventor(s):  
Anoop Gupta  
David M. Barger

ATTORNEY'S DOCKET NO. MS1-302US

1 A portion of the disclosure of this patent document contains material which  
2 is subject to copyright protection. The copyright owner has no objection to the  
3 facsimile reproduction by anyone of the patent document or the patent disclosure,  
4 as it appears in the Patent and Trademark Office patent file or records, but  
5 otherwise reserves all copyright rights whatsoever.

## 6 7 **RELATED APPLICATIONS**

8 This application claims priority to U.S. Provisional Application No.  
9 60/100,452, filed September 15, 1998, entitled "Annotations for Streaming Video  
10 on the Web: System Design and Usage", to Anoop Gupta and David M. Barger.

## 11 12 **TECHNICAL FIELD**

13 This invention relates to networked client/server systems and to methods of  
14 delivering and rendering multimedia content in such systems. More particularly,  
15 the invention relates to systems and methods of selecting and providing such  
16 content.

## 17 18 **BACKGROUND OF THE INVENTION**

19 The advent of computers and their continued technological advancement  
20 has revolutionized the manner in which people work and live. An example of  
21 such is in the education field, wherein educational presentations (such as college  
22 lectures, workplace training sessions, etc.) can be provided to a computer user as  
23 multimedia data (e.g., video, audio, text, and/or animation data). Today, such  
24 presentations are primarily video and audio, but a richer, broader digital media era  
25 is emerging. Educational multimedia presentations provide many benefits, such as

1 allowing the presentation data to be created at a single time yet be presented to  
2 different users at different times and in different locations throughout the world.

3 These multimedia presentations are provided to a user as synchronized  
4 media. Synchronized media means multiple media objects that share a common  
5 timeline. Video and audio are examples of synchronized media—each is a  
6 separate data stream with its own data structure, but the two data streams are  
7 played back in synchronization with each other. Virtually any media type can  
8 have a timeline. For example, an image object can change like an animated .gif  
9 file, text can change and move, and animation and digital effects can happen over  
10 time. This concept of synchronizing multiple media types is gaining greater  
11 meaning and currency with the emergence of more sophisticated media  
12 composition frameworks implied by MPEG-4, Dynamic HTML, and other media  
13 playback environments.

14 The term “streaming” is used to indicate that the data representing the  
15 various media types is provided over a network to a client computer on a real-  
16 time, as-needed basis, rather than being pre-delivered in its entirety before  
17 playback. Thus, the client computer renders streaming data as it is received from a  
18 network server, rather than waiting for an entire “file” to be delivered.

19 Multimedia presentations may also include “annotations” relating to the  
20 multimedia presentation. An annotation is data (e.g., audio, text, video, etc.) that  
21 corresponds to a multimedia presentation. Annotations can be added by anyone  
22 with appropriate access rights to the annotation system (e.g., the lecturer/trainer or  
23 any of the students/trainees). These annotations typically correspond to a  
24 particular temporal location in the multimedia presentation and can provide a  
25 replacement for much of the “in-person” interaction and “classroom discussion”

1 that is lost when the presentation is not made "in-person" or "live". As part of an  
2 annotation, a student can comment on a particular point, to which another student  
3 (or lecturer) can respond in a subsequent annotation. This process can continue,  
4 allowing a "classroom discussion" to occur via these annotations. Additionally,  
5 some systems allow a user to select a particular one of these annotations and begin  
6 playback of the presentation starting at approximately the point in the presentation  
7 to which the annotation corresponds.

8 However, current systems typically allow a user to select multimedia  
9 playback based only on individual annotations. This limitation provides a  
10 cumbersome process for the user, as he or she may wish to view several different  
11 portions of the presentation corresponding to several different annotations. Using  
12 current systems, the user would be required to undergo the painstaking process of  
13 selecting a first annotation, viewing/listening to the multimedia presentation  
14 corresponding to the first annotation, selecting a second annotation,  
15 viewing/listening to the multimedia presentation corresponding to the second  
16 annotation, selecting a third annotation, viewing/listening to the multimedia  
17 presentation corresponding to the third annotation, and so on through several  
18 annotations.

19 The invention described below addresses this and other disadvantages of  
20 annotations, providing a way to improve multimedia presentation using  
21 annotations.

## 22 23 SUMMARY OF THE INVENTION

24 Annotations correspond to media segments of one or more multimedia  
25 streams. A playlist generation interface is presented to the user in the form of

1 annotation titles or summaries for a group of annotations. This group of  
2 annotations corresponds to the media segments that are part of a playlist. The  
3 playlist can then be altered by the user to suit his or her desires or needs by  
4 interacting with the annotation title/summary interface. The media segments of  
5 the playlist can then be presented to the user in a seamless, contiguous manner.

6 According to one aspect of the invention, the ordering of the annotation  
7 titles/summaries can be altered by the user, resulting in a corresponding change in  
8 order of presentation of the media segments. The ordering of the annotation  
9 titles/summaries can be changed by moving the titles or summaries in a drag and  
10 drop manner.

11 According to another aspect of the invention, the media segments of the  
12 playlist can themselves be stored as an additional multimedia stream. This  
13 additional multimedia stream can then be annotated in the same manner as other  
14 multimedia streams.

## 15 **BRIEF DESCRIPTION OF THE DRAWINGS**

16  
17 Fig. 1 shows a client/server network system and environment in accordance  
18 with one embodiment of the invention.

19 Fig. 2 shows a general example of a computer that can be used as a client or  
20 server in accordance with the invention.

21 Fig. 3 is a block diagram illustrating an annotation server and a client  
22 computer in more detail in accordance with one embodiment of the invention.

23 Fig. 4 is a block diagram illustrating the structure for an annotation  
24 according to one embodiment of the invention.

25 Fig. 5 is a block diagram illustrating exemplary annotation collections.

1 Fig. 6 illustrates an annotation toolbar in accordance with one embodiment  
2 of the invention.

3 Fig. 7 illustrates an "add new annotation" dialog box in accordance with  
4 one embodiment of the invention.

5 Fig. 8 illustrates a "query annotations" dialog box in accordance with one  
6 embodiment of the invention.

7 Fig. 9 illustrates a "view annotations" dialog box in accordance with one  
8 embodiment of the invention.

9 Fig. 10 is a diagrammatic illustration of a graphical user interface window  
10 displaying annotations and corresponding media segments concurrently in  
11 accordance with one embodiment of the invention.

12 Fig. 11 illustrates methodological aspects of one embodiment of the  
13 invention in retrieving and presenting annotations and media segments to a user.

## 14 **DETAILED DESCRIPTION**

### 15 **General Network Structure**

16 Fig. 1 shows a client/server network system and environment in accordance  
17 with one embodiment of the invention. Generally, the system includes multiple  
18 network server computers 10, 11, 12, and 13, and multiple ( $n$ ) network client  
19 computers 15. The computers communicate with each other over a data  
20 communications network. The communications network in Fig. 1 comprises a  
21 public network 16 such as the Internet. The data communications network might  
22 also include, either in addition to or in place of the Internet, local-area networks  
23 and/or private wide-area networks.  
24  
25

1 Streaming media server computer 11 has access to streaming media content  
2 in the form of different media streams. These media streams can be individual  
3 media streams (e.g., audio, video, graphical, etc.), or alternatively can be  
4 composite media streams including two or more of such individual streams. Some  
5 media streams might be stored as files in a database or other file storage system,  
6 while other media streams might be supplied to the server on a "live" basis from  
7 other data source components through dedicated communications channels or  
8 through the Internet itself.

9 There are various standards for streaming media content and composite  
10 media streams. The "Advanced Streaming Format" (ASF) is an example of such a  
11 standard, including both accepted versions of the standard and proposed standards  
12 for future adoption. ASF specifies the way in which multimedia content is stored,  
13 streamed, and presented by the tools, servers, and clients of various multimedia  
14 vendors. Further details about ASF are available from Microsoft Corporation of  
15 Redmond, Washington.

16 Annotation server 10 controls the storage of annotations and their provision  
17 to client computers 15. The annotation server 10 manages the annotation meta  
18 data store 18 and the annotation content store 17. The annotation server 10  
19 communicates with the client computers 15 via any of a wide variety of known  
20 protocols, such as the Hypertext Transfer Protocol (HTTP). The annotation server  
21 10 can receive and provide annotations via direct contact with a client computer  
22 15, or alternatively via electronic mail (email) via email server 13. The annotation  
23 server 10 similarly communicates with the email server 13 via any of a wide  
24 variety of known protocols, such as the Simple Mail Transfer Protocol (SMTP).



1 The annotations managed by annotation server 10 correspond to the  
2 streaming media available from media server computer 11. In the discussions to  
3 follow, the annotations are discussed as corresponding to streaming media.  
4 However, it should be noted that the annotations can similarly correspond to "pre-  
5 delivered" rather than streaming media, such as media previously stored at the  
6 client computers 15 via the network 16, via removable magnetic or optical disks,  
7 etc.

8 When a user of a client computer 15 accesses a web page containing  
9 streaming media, a conventional web browser of the client computer 15 contacts  
10 the web server 12 to get the Hypertext Markup Language (HTML) page, the media  
11 server 11 to get the streaming data, and the annotation server 10 to get any  
12 annotations associated with that media. When a user of a client computer 15  
13 desires to add or retrieve annotations, the client computer 15 contacts the  
14 annotation server 10 to perform the desired addition/retrieval.

### 15 16 **Exemplary Computer Environment**

17 In the discussion below, the invention will be described in the general  
18 context of computer-executable instructions, such as program modules, being  
19 executed by one or more conventional personal computers. Generally, program  
20 modules include routines, programs, objects, components, data structures, etc. that  
21 perform particular tasks or implement particular abstract data types. Moreover,  
22 those skilled in the art will appreciate that the invention may be practiced with  
23 other computer system configurations, including hand-held devices,  
24 multiprocessor systems, microprocessor-based or programmable consumer  
25 electronics, network PCs, minicomputers, mainframe computers, and the like. In a

1 distributed computer environment, program modules may be located in both local  
2 and remote memory storage devices.

3 Fig. 2 shows a general example of a computer 20 that can be used as a  
4 client or server in accordance with the invention. Computer 20 is shown as an  
5 example of a computer that can perform the functions of any of server computers  
6 10-13 or a client computer 15 of Figure 1.

7 Computer 20 includes one or more processors or processing units 21, a  
8 system memory 22, and a bus 23 that couples various system components  
9 including the system memory 22 to processors 21.

10 The bus 23 represents one or more of any of several types of bus structures,  
11 including a memory bus or memory controller, a peripheral bus, an accelerated  
12 graphics port, and a processor or local bus using any of a variety of bus  
13 architectures. The system memory includes read only memory (ROM) 24 and  
14 random access memory (RAM) 25. A basic input/output system (BIOS) 26,  
15 containing the basic routines that help to transfer information between elements  
16 within computer 20, such as during start-up, is stored in ROM 24. Computer 20  
17 further includes a hard disk drive 27 for reading from and writing to a hard disk,  
18 not shown, a magnetic disk drive 28 for reading from and writing to a removable  
19 magnetic disk 29, and an optical disk drive 30 for reading from or writing to a  
20 removable optical disk 31 such as a CD ROM or other optical media. The hard  
21 disk drive 27, magnetic disk drive 28, and optical disk drive 30 are connected to  
22 the system bus 23 by an SCSI interface 32 or some other appropriate interface.  
23 The drives and their associated computer-readable media provide nonvolatile  
24 storage of computer readable instructions, data structures, program modules and  
25 other data for computer 20. Although the exemplary environment described

1 herein employs a hard disk, a removable magnetic disk 29 and a removable optical  
2 disk 31, it should be appreciated by those skilled in the art that other types of  
3 computer readable media which can store data that is accessible by a computer,  
4 such as magnetic cassettes, flash memory cards, digital video disks, random access  
5 memories (RAMs) read only memories (ROM), and the like, may also be used in  
6 the exemplary operating environment.

7 A number of program modules may be stored on the hard disk, magnetic  
8 disk 29, optical disk 31, ROM 24, or RAM 25, including an operating system 35,  
9 one or more application programs 36, other program modules 37, and program  
10 data 38. A user may enter commands and information into computer 20 through  
11 input devices such as keyboard 40 and pointing device 42. Other input devices  
12 (not shown) may include a microphone, joystick, game pad, satellite dish, scanner,  
13 or the like. These and other input devices are connected to the processing unit 21  
14 through an interface 46 that is coupled to the system bus. A monitor 47 or other  
15 type of display device is also connected to the system bus 23 via an interface, such  
16 as a video adapter 48. In addition to the monitor, personal computers typically  
17 include other peripheral output devices (not shown) such as speakers and printers.

18 Computer 20 operates in a networked environment using logical  
19 connections to one or more remote computers, such as a remote computer 49. The  
20 remote computer 49 may be another personal computer, a server, a router, a  
21 network PC, a peer device or other common network node, and typically includes  
22 many or all of the elements described above relative to computer 20, although only  
23 a memory storage device 50 has been illustrated in Fig. 2. The logical connections  
24 depicted in Fig. 2 include a local area network (LAN) 51 and a wide area network  
25 (WAN) 52. Such networking environments are commonplace in offices,

1 enterprise-wide computer networks, intranets, and the Internet. In the described  
2 embodiment of the invention, remote computer 49 executes an Internet Web  
3 browser program such as the "Internet Explorer" Web browser manufactured and  
4 distributed by Microsoft Corporation of Redmond, Washington.

5 When used in a LAN networking environment, computer 20 is connected to  
6 the local network 51 through a network interface or adapter 53. When used in a  
7 WAN networking environment, computer 20 typically includes a modem 54 or  
8 other means for establishing communications over the wide area network 52, such  
9 as the Internet. The modem 54, which may be internal or external, is connected to  
10 the system bus 23 via a serial port interface 33. In a networked environment,  
11 program modules depicted relative to the personal computer 20, or portions  
12 thereof, may be stored in the remote memory storage device. It will be  
13 appreciated that the network connections shown are exemplary and other means of  
14 establishing a communications link between the computers may be used.

15 Generally, the data processors of computer 20 are programmed by means of  
16 instructions stored at different times in the various computer-readable storage  
17 media of the computer. Programs and operating systems are typically distributed,  
18 for example, on floppy disks or CD-ROMs. From there, they are installed or  
19 loaded into the secondary memory of a computer. At execution, they are loaded at  
20 least partially into the computer's primary electronic memory. The invention  
21 described herein includes these and other various types of computer-readable  
22 storage media when such media contain instructions or programs for implementing  
23 the steps described below in conjunction with a microprocessor or other data  
24 processor. The invention also includes the computer itself when programmed  
25 according to the methods and techniques described below. Furthermore, certain

1 sub-components of the computer may be programmed to perform the functions  
2 and steps described below. The invention includes such sub-components when  
3 they are programmed as described. In addition, the invention described herein  
4 includes data structures, described below, as embodied on various types of  
5 memory media.

6 For purposes of illustration, programs and other executable program  
7 components such as the operating system are illustrated herein as discrete blocks,  
8 although it is recognized that such programs and components reside at various  
9 times in different storage components of the computer, and are executed by the  
10 data processor(s) of the computer.

### 11 12 Client/Server Relationship

13 Fig. 3 illustrates an annotation server and a client computer in more detail.  
14 As noted above, generally, commands are formulated at client computer 15 and  
15 forwarded to annotation server 10 via HTTP requests. In the illustrated  
16 embodiment of Fig. 3, communication between client 15 and server 10 is  
17 performed via HTTP, using commands encoded as Uniform Resource Locators  
18 (URLs) and data formatted as object linking and embedding (OLE) structured  
19 storage documents, or alternatively using Extensible Markup Language (XML).

20 Client 15 includes an HTTP services (HttpSvcs) module 152, which  
21 manages communication with server 10, and an annotation back end (ABE)  
22 module 151, which translates user actions into commands destined for server 10.  
23 A user interface (MMA) module 150 provides the user interface (UI) for a user to  
24 add and select different annotations, and be presented with the annotations.  
25

1 According to one implementation, the user interface module 150 supports ActiveX  
2 controls that display an annotation interface for streaming video on the Web.

3 Client 15 also includes a web browser module 153, which provides a  
4 conventional web browsing interface and capabilities for the user to access various  
5 servers via network 16 of Fig. 1. Web browser 153 also provides the interface for  
6 a user to be presented with media streams. In addition to the use of playlists  
7 discussed below, the user can select which one of different versions of multimedia  
8 content he or she wishes to receive from media server 11 of Fig. 1. This selection  
9 can be direct (e.g., entry of a particular URL or selection of a "low resolution"  
10 option), or indirect (e.g., entry of a particular desired playback duration or an  
11 indication of system capabilities, such as "slow system" or "fast system").  
12 Alternatively, other media presentation interfaces could be used.

13 Annotation server 10 includes the Multimedia Annotation Web Server  
14 (MAWS) module 130, which is an Internet Services Application Programming  
15 Interface (ISAPI) plug-in for Internet Information Server (IIS) module 135.  
16 Together, these two modules provide the web server functionality of annotation  
17 server 10. Annotation server 10 also includes an HTTP Services module 131  
18 which manages communication with client 15. In addition, annotation server 10  
19 utilizes The Windows Messaging Subsystem 134 to facilitate communication with  
20 email server 13 of Fig. 1, and an email reply server 133 for processing incoming  
21 email received from email server 13.

22 Annotation server 10 further includes an annotation back end (ABE)  
23 module 132, which contains functionality for accessing annotation stores 17 and  
24 18, for composing outgoing email based on annotation data, and for processing  
25 incoming email. Incoming email is received and passed to the ABE module 132

1 by the Email Reply Server 133. Annotation content authored at client 15, using  
2 user interface 150, is received by ABE 132 and maintained in annotation content  
3 store 17. Received meta data (control information) corresponding to the  
4 annotation content is maintained in annotation meta data store 18. The annotation  
5 content and meta data can be stored in any of a variety of conventional manners,  
6 such as in SQL relational databases (e.g., using Microsoft "SQL Server" version  
7 7.0, available from Microsoft Corporation). Annotation server 10 is illustrated in  
8 Fig. 3 as maintaining the annotation content and associated control information  
9 (meta data) separately in two different stores. Alternatively, all of the annotation  
10 data (content and meta information) can be stored together in a single store, or  
11 content may be stored by another distinct storage system on the network 16 of Fig.  
12 1, such as a file system, media server, email server, or other data store.

13 ABE 132 of annotation server 10 also manages the interactive generation  
14 and presentation of streaming media data from server computer 11 of Fig. 1 using  
15 "playlists". A "playlist" is a listing of one or more multimedia segments to be  
16 retrieved and presented in a given order. Each of the multimedia segments in the  
17 playlist is defined by a source identifier, a start time, and an end time. The source  
18 identifier identifies which media stream the segment is part of, the start time  
19 identifies the temporal location within the media stream where the segment begins,  
20 and the end time identifies the temporal location within the media stream where  
21 the segment ends.

22 ABE 132 allows playlists to be generated interactively based on annotations  
23 maintained in annotation stores 17 and 18. ABE 132 provides a user at client 15  
24 with multiple possible annotation identifiers (e.g., titles or summaries) from which  
25 the user can select those of interest to him or her. Based on the selected



1 annotations, ABE 132 coordinates provision of the associated media segments to  
2 the user. ABE 132 can directly communicate with video server computer 11 to  
3 identify which segments are to be provided, or alternatively can provide the  
4 appropriate information to the browser of client computer 15, which in turn can  
5 request the media segments from server computer 11.

6 Fig. 4 shows an exemplary structure for an annotation entry 180 that is  
7 maintained by annotation server 10 in annotation meta data store 18 of Fig. 3. In  
8 the illustrated embodiment, an annotation entry 180 includes an author field 182, a  
9 time range field 184, a time units field 186, a creation time field 188, a title field  
10 190, a content field 192, an identifier field 194, a related annotation identifier field  
11 196, a set identifier(s) field 198, a media content identifier field 200, an arbitrary  
12 number of user-defined property fields 202, and a sequence number 204. Each of  
13 fields 182-204 is a collection of data which define a particular characteristic of  
14 annotation entry 180. Annotation entry 180 is maintained by annotation server 10  
15 of Fig. 3 in annotation meta data store 18. Content field 192, as discussed in more  
16 detail below, includes a pointer to (or other identifier of) the annotation content,  
17 which in turn is stored in annotation content store 17.

18 Author field 182 contains data identifying the user who created annotation  
19 entry 180 and who is therefore the author of the annotation. The author is  
20 identified by ABE 151 of Fig. 3 based on the user logged into client 15 at the time  
21 the annotation is created.

22 Time range field 184 contains data representing "begin" and "end" times  
23 defining a segment of media timeline to which annotation entry 180 is associated.  
24 Time units field 186 contains data representing the units of time represented in  
25 time range field 184. Together, time range field 184 and time units field 186



1 identify the relative time range of the annotation represented by annotation entry  
2 180. This relative time range corresponds to a particular segment of the media  
3 stream to which annotation entry 180 is associated. The begin and end times for  
4 the annotation are provided by the user via interface 150 of Fig. 3, or alternatively  
5 can be automatically or implicitly derived using a variety of audio and video  
6 signal processing techniques, such as sentence detection in audio streams or video  
7 object tracking.

8 It should be noted that the time ranges for different annotations can overlap.  
9 Thus, for example, a first annotation may correspond to a segment ranging  
10 between the first and fourth minutes of media content, a second annotation may  
11 correspond to a segment ranging between the second and seventh minutes of the  
12 media content, and a third annotation may correspond to a segment ranging  
13 between the second and third minutes of the media content.

14 Alternatively, rather than using the presentation timeline of the media  
15 content, different media characteristics can be used to associate the annotation  
16 with a particular segment(s) of the media content. For example, annotations could  
17 be associated with (or “anchored” on) specific objects in the video content, or  
18 specific events in the audio content.

19 Creation time field 188 contains data specifying the date and time at which  
20 annotation entry 180 is created. It should be noted that the time of creation of  
21 annotation entry 180 is absolute and is not relative to the video or audio content of  
22 the media stream to which annotation entry 180 is associated. Accordingly, a user  
23 can specify that annotations which are particularly old, e.g., created more than two  
24 weeks earlier, are not to be displayed. ABE 132 of Fig. 3 stores the creation time  
25 and date when the annotation is created.

1 Title field 190 contains data representing a title by which the annotation  
2 represented by annotation entry 180 is identified. The title is generally determined  
3 by the user and the user enters the data representing the title using conventional  
4 and well known user interface techniques. The data can be as simple as ASCII  
5 text or as complex as HTML code which can include text having different fonts  
6 and type styles, graphics including wallpaper, motion video images, audio, and  
7 links to other multimedia documents.

8 Content field 192 contains data representing the substantive content of the  
9 annotation as authored by the user. The actual data can be stored in content field  
10 192, or alternatively content field 192 may store a pointer to (or other indicator of)  
11 the content that is stored separately from the entry 180 itself. In the illustrated  
12 example, content field 192 includes a pointer to (or other identifier of) the  
13 annotation content, which in turn is stored in annotation content store 17. The user  
14 enters the data representing the content using conventional and well known user  
15 interface techniques. The content added by the user in creating annotation entry  
16 180 can include any one or more of text, graphics, video, audio, etc. or links  
17 thereto. In essence, content field 192 contains data representing the substantive  
18 content the user wishes to include with the presentation of the corresponding  
19 media stream at the relative time range represented by time range field 184 and  
20 time units field 186.

21 Annotation identifier field 194 stores data that uniquely identifies  
22 annotation entry 180, while related annotation identifier field 196 stores data that  
23 uniquely identifies a related annotation. Annotation identifier field 194 can be  
24 used by other annotation entries to associate such other annotation entries with  
25 annotation entry 180. In this way, threads of discussion can develop in which a

1 second annotation responds to a first annotation, a third annotation responds to the  
2 second annotation and so on. By way of example, an identifier of the first  
3 annotation would be stored in related annotation identifier field 196 of the second  
4 annotation, an identifier of the second annotation would be stored in related  
5 annotation identifier field 196 of the third annotation, and so on.

6 Set identifier(s) field 198 stores data that identifies a particular one or more  
7 sets to which annotation entry 180 belongs. A media stream can have multiple  
8 sets of annotations, sets can span multiple media content, and a particular  
9 annotation can correspond to one or more of these sets. Which set(s) an  
10 annotation belongs to is identified by the author of the annotation. By way of  
11 example, a media stream corresponding to a lecture may include the following  
12 sets: "instructor's comments", "assistant's comments", "audio comments", "text  
13 comments", "student questions", and each student's personal comments.

14 Media content identifier field 200 contains data that uniquely identifies  
15 particular multimedia content as the content to which annotation entry 180  
16 corresponds. Media content identifier 200 can identify a single media stream  
17 (either an individual stream or a composite stream), or alternatively identify  
18 multiple different streams that are different versions of the same media content.  
19 Media content identifier 200 can identify media versions in a variety of different  
20 manners. According to one embodiment, the data represents a real-time transport  
21 protocol (RTP) address of the different media streams. An RTP address is a type  
22 of uniform resource locator (URL) by which multimedia documents can be  
23 identified in a network. According to an alternate embodiment, a unique identifier  
24 is assigned to the content rather than to the individual media streams. According  
25 to another alternate embodiment, a different unique identifier of the media streams

1 could be created by annotation server 10 of Fig. 3 and assigned to the media  
2 streams. Such a unique identifier would also be used by streaming media server  
3 11 of Fig. 1 to identify the media streams. According to another alternate  
4 embodiment, a uniform resource name (URN) such as those described by K.  
5 Sollins and L. Mosinter in "Functional Requirements for Uniform Resource  
6 Names," IETF RFC 1733 (December 1994) could be used to identify the media  
7 stream.

8 User-defined property fields 202 are one or more user-definable fields that  
9 allow users (or user interface designers) to customize the annotation system.  
10 Examples of such additional property fields include a "reference URL" property  
11 which contains the URL of a web page used as reference material for the content  
12 of the annotation; a "help URL" property containing the URL of a help page which  
13 can be accessed concerning the content of the annotation; a "view script" property  
14 containing JavaScript which is to be executed whenever the annotation is viewed;  
15 a "display type" property, which gives the client user interface information about  
16 how the annotation is to be displayed; etc.

17 Sequence number 204 allows a user to define (via user interface 150 of Fig.  
18 3) a custom ordering for the display of annotation identifiers, as discussed in more  
19 detail below. Sequence number 204 stores the relative position of the annotations  
20 with respect to one another in the custom ordering, allowing the custom ordering  
21 to be saved for future used. In the illustrated example, annotation entry 180 stores  
22 a single sequence number. Alternatively, multiple sequence numbers 204 may be  
23 included in annotation entry 180 each corresponding to a different custom  
24 ordering, or a different annotation set, or a different user, etc.  
25

1 Fig. 5 illustrates exemplary implicit annotation collections for annotations  
2 maintained by annotation server 10 of Fig. 3. A collection of annotations refers to  
3 annotation entries 180 of Fig. 4 that correspond to the same media stream(s),  
4 based on the media content identifier 200. Annotation entries 180 can be viewed  
5 conceptually as part of the same annotation collection if they have the same media  
6 content identifiers 200, even though the annotation entries may not be physically  
7 stored together by annotation server 10.

8 Annotation database 206 includes two annotation collections 208 and 210.  
9 Annotation server 10 dynamically adds, deletes, and modifies annotation entries in  
10 annotation database 206 based on commands from client 15. Annotation entries  
11 can be created and added to annotation database 206 at any time a user cares to  
12 comment upon the content of the stream (or another annotation) in the form of an  
13 annotation. Annotation server 10 forms an annotation entry from identification  
14 data, content data, title data, and author data of an "add annotation" request  
15 received from the client's ABE 151 (Fig. 3), and adds the annotation entry to  
16 annotation database 206.

17 Annotation database 206 includes a fields 212, 214, and 216 that specify  
18 common characteristics of all annotation entries of database 206 or an annotation  
19 collection 208 or 210. Alternatively, fields 212-216 can be included redundantly  
20 in each annotation entry 180.

21 Creator field 212 contains data identifying the user who was responsible for  
22 creating annotation database 206.

23 RTP address fields 214 and 216 contain data representing an RTP address  
24 of the media stream (e.g., the RTP address of the stream identified in media  
25 content identifier 200 of Fig. 5) for the annotation collection. An RTP address

1 provides an alternative mechanism, in addition to the data in identifier field 200,  
2 for associating the media stream with annotation entries 180. In alternative  
3 embodiments, RTP address fields 214 and 216 need not be included, particularly  
4 embodiments in which media content identifier 200 contains the RTP address of  
5 the media stream.

## 6 7 **User Interface**

8 An annotation can be created by a user of any of the client computers 15 of  
9 Fig. 1. As discussed above with reference to Fig. 3, client 15 includes an interface  
10 module 150 that presents an interface to a user (e.g., a graphical user interface),  
11 allowing a user to make requests of annotation server 10. In the illustrated  
12 embodiment, a user can access annotation server 10 via an annotation toolbar  
13 provided by interface 150.

14 Fig. 6 illustrates an annotation toolbar in accordance with one embodiment  
15 of the invention. Annotation toolbar 240 includes various identifying information  
16 and user-selectable options 242-254.

17 Selection of an exit or "X" button 242 causes interface 150 to terminate  
18 display of the toolbar 240. A server identifier 244 identifies the annotation server  
19 with which client 15 is currently configured to communicate (annotation server 10  
20 of Fig. 1. in the illustrated embodiment).

21 Selection of a connection button 246 causes ABE 151 of Fig. 3 to establish  
22 a connection with the annotation server identified by identifier 244. Selection of a  
23 query button 248 causes interface module 150 to open a "query" dialog box, from  
24 which a user can search for particular annotations. Selection of an add button 250  
25

1 causes interface module 150 to open an "add new annotation" dialog box, from  
2 which a user can create a new annotation.

3 Selection of a show annotations button 252 causes interface module 150 to  
4 open a "view annotations" dialog box, from which a user can select particular  
5 annotations for presentation.

6 Selection of a preferences button 254 causes interface 150 of Fig. 3 to open  
7 a "preferences" dialog box, from which a user can specify various UI preferences,  
8 such as an automatic server query refresh interval, or default query criteria values  
9 to be persisted between sessions.

### 10 11 Annotation Creation

12 Fig. 7 shows an "add new annotation" dialog box 260 that results from user  
13 selection of add button 250 of Fig. 6 to create a new annotation. Interface 150  
14 assumes that the current media stream being presented to the user is the media  
15 stream to which this new annotation will be associated. The media stream to  
16 which an annotation is associated is referred to as the "target" of the annotation.  
17 An identifier of this stream is displayed in a target specification area 262 of dialog  
18 box 260. Alternatively, a user could change the target of the annotation, such as  
19 by typing in a new identifier in target area 262, or by selection of a "browse"  
20 button (not shown) that allows the user to browse through different directories of  
21 media streams.

22 A time strip 264 is also provided as part of dialog box 260. Time strip 264  
23 represents the entire presentation time of the corresponding media stream. A  
24 thumb 265 that moves within time strip 264 indicates a particular temporal  
25 position within the media stream. The annotation being created via dialog box 260



1 has a begin time and an end time, which together define a particular segment of  
2 the media stream. When "from" button 268 is selected, thumb 265 represents the  
3 begin time for the segment relative to the media stream. When "to" button 271 is  
4 selected, thumb 265 represents the end time for the segment relative to the media  
5 stream. Alternatively, two different thumbs could be displayed, one for the begin  
6 time and one for the end time. The begin and end times are also displayed in an  
7 hours/minutes/seconds format in boxes 266 and 270, respectively.

8 Thumb 265 can be moved along time strip 264 in any of a variety of  
9 conventional manners. For example, a user can depress a button of a mouse (or  
10 other cursor control device) while a pointer is "on top" of thumb 265 and move the  
11 pointer along time strip 264, causing thumb 265 to move along with the pointer.  
12 The appropriate begin or end time is then set when the mouse button is released.  
13 Alternatively, the begin and end times can be set by entering (e.g., via an  
14 alphanumeric keyboard) particular times in boxes 266 and 270.

15 Dialog box 260 also includes a "play" button 274. Selection of play button  
16 274 causes interface module 150 of Fig. 3 to forward a segment specification to  
17 web browser 153 of client 15. The segment specification includes the target  
18 identifier from target display 262 and the begin and end times from boxes 266 and  
19 270, respectively. Upon receipt of the segment specification from interface  
20 module 150, the browser communicates with media server 11 and requests the  
21 identified media segment using conventional HTTP requests. In response, media  
22 server 11 streams the media segment to client 15 for presentation to the user. This  
23 presentation allows, for example, the user to verify the portion of the media stream  
24 to which his or her annotation will correspond.



1 Dialog box 260 also includes an annotation set identifier 272, an email field  
2 275, and a summary 276. Annotation set identifier 272 allows the user to identify  
3 a named set to which the new annotation will belong. This set can be a previously  
4 defined set, or a new set being created by the user. Selection of the particular set  
5 can be made from a drop-down menu activated by selection of a button 273, or  
6 alternatively can be directly input by the user (e.g., typed in using an alphanumeric  
7 keyboard). According to one embodiment of the invention, annotation server 10  
8 of Fig. 3 supports read and write access controls, allowing the creator of the set to  
9 identify which users are able to read and/or write to the annotation set. In this  
10 embodiment, only those sets for which the user has write access can be entered as  
11 set identifier 272.

12 Email identifier 275 allows the user to input the email address of a recipient  
13 of the annotation. When an email address is included, the newly created  
14 annotation is electronically mailed to the recipient indicated in identifier 275 in  
15 addition to being added to the annotation database. Furthermore, the recipient of  
16 the electronic mail message can reply to the message to create an additional  
17 annotation. To enable this, the original email message is configured with  
18 annotation server 10 as the sender. Because of this, a "reply to sender" request  
19 from the recipient will cause an email reply to be sent to annotation server 10.  
20 Upon receipt of such an electronic mail message reply, annotation server 10  
21 creates a new annotation and uses the reply message content as the content of the  
22 new annotation. This new annotation identifies, as a related annotation, the  
23 original annotation that was created when the original mail message was sent by  
24 annotation server 10. In the illustrated embodiment, this related annotation  
25 identifier is stored in field 196 of Fig 4.

Summary 276 allows the user to provide a short summary or title of the annotation content. Although the summary is illustrated as being text, it could include any of a wide variety of characters, alphanumerics, graphics, etc. In the illustrated embodiment, summary 276 is stored in the title field 190 of the annotation entry of Fig. 4.

Dialog box 260 further includes radio buttons 280 and 282, which allow an annotation to be created as text and/or audio. Although not shown, other types of annotations could also be accommodated, such as graphics, HTML documents, etc. Input controls 278 are also provided as part of dialog box. The illustrated controls are enabled when the annotation includes audio data. Input controls 278 include conventional audio control buttons such as fast forward, rewind, play, pause, stop and record. Additionally, an audio display bar 279 can be included to provide visual progress feedback when the audio is playing or recording.

The exact nature of input controls 278 is dependent on the type of annotation content being provided. In the case of text content, input controls 278 may simply include a box into which text can be input by the user via an alphanumeric keyboard. Additionally, a keyboard layout may also be provided to the user, allowing him or her to “point and click” using a mouse and pointer to select particular characters for entry.

### **Annotation and Media Segment Retrieval**

Fig. 8 shows a “query annotations” dialog box 330 that results from a user selecting query button 248 of Fig. 6. Many of the options presented to the user in dialog box 330 are similar to those presented in the “add new annotation” dialog

1 box 260 of Fig. 7, however, those in dialog box 330 are used as search criteria  
2 rather than data for a new annotation.

3 Dialog box 330 includes a target display 332 that contains an identifier of  
4 the target stream. This identifier can be input in any of a variety of manners, such  
5 as by typing in a new identifier in target display 332, or by selection of a "browse"  
6 button (not shown) that allows the user to browse through different directories of  
7 media streams. In the illustrated embodiment, the identifier is an URL. However,  
8 alternate embodiments can use different identifier formats.

9 Dialog box 330 also includes target information 334, which includes a time  
10 strip, thumb, "from" button, "to" button, "play" button, and begin and end times,  
11 which are analogous to the time strip, thumb, "from" button, "to" button, "play"  
12 button, begin and end times of dialog box 260 of Fig. 7. The begin and end times  
13 in target information 334 limit the query for annotations to only those annotations  
14 having a time range that corresponds to at least part of the media segment between  
15 the begin and end times of target information 334.

16 Dialog box 330 also includes an annotation set list 336. Annotation set list  
17 336 includes a listing of the various sets that correspond to the target media  
18 stream. According to one implementation, only those sets for which an annotation  
19 has been created are displayed in set list 336. According to one embodiment of  
20 the invention, annotation server 10 of Fig. 3 supports read and write security,  
21 allowing the creator of the set to identify which users are able to read and/or write  
22 to the annotation set. In this embodiment, only those sets for which the user has  
23 read access are displayed in set list 336.

24 A user can select sets from annotation set list 336 in a variety of manners.  
25 For example, using a mouse and pointer to "click" on a set in list 336, which

1 highlights the set to provide feedback to the user that the set has been selected.  
2 Clicking on the selected set again de-selects the set (leaving it no longer  
3 highlighted). Additionally, a "select all" button 338 allows a user to select all sets  
4 in set list 336, while a "deselect all" button 340 allows a user to de-select all sets  
5 in set list 336.

6 In the illustrated embodiment, the sets displayed as part of annotation set  
7 list 336 contain annotations which correspond to the target identifier in target  
8 display 332. However, in alternate embodiments the sets in selection list 338 need  
9 not necessarily contain annotations which correspond to the target identifier in  
10 target display 332. Interface module 150 allows a user to select different target  
11 streams during the querying process. Thus, a user may identify a first target  
12 stream and select one or more sets to query annotations from for the first target  
13 stream, and then identify a second target stream and select one or more sets to  
14 query annotations from for the second target stream.

15 Additional search criteria can also be input by the user. As illustrated, a  
16 particular creation date and time identifier 342 can be input, along with a relation  
17 344 (e.g., "after" or "before"). Similarly, particular words, phrases, characters,  
18 graphics, etc. that must appear in the summary can be input in a summary keyword  
19 search identifier 346. A maximum number of annotations to retrieve in response  
20 to the query can also be included as a max identifier 348. Furthermore, the query  
21 can be limited to only annotations that correspond to the target identifier in target  
22 display 332 by selecting check box 360.

23 A level of detail 350 to retrieve can also be selected by the user. Examples  
24 of different levels that could be retrieved include the "full level" (that is, all  
25 content of the annotation), or a "deferred download" where only an identifier of

1 the annotations (e.g., a summary or title) is downloaded. In the illustrated  
2 example, selection of checkbox 354 selects the deferred download level, whereas  
3 if checkbox 354 is not selected then the full level of detail is implicitly selected.

4 A server identifier 356 identifies the annotation server with which client 15  
5 is currently configured to communicate. Different annotation servers can be  
6 selected by the user by inputting the appropriate identifier as server identifier 356.  
7 This input can be provided in any of a variety of manners, such as by typing in a  
8 new identifier in server identifier 356 or by selection of a "browse" button (not  
9 shown) that allows the user to browse through different directories of annotation  
10 servers.

11 A user can request automatic display of the retrieved annotations by  
12 selecting a "display retrieved annotations" checkbox 358. Selection of  
13 "advanced" button 362 reduces the number of options available to the user,  
14 simplifying dialog box 330. For example, the simplified dialog box may not  
15 include fields 342, 344, 348, 346, 350, 332, 334, or 336.

16 The user can then complete the query process by selecting a query button  
17 364. Upon selection of the query button 364, interface 150 closes the query dialog  
18 box 330 and forwards the search criteria to annotation server 10. Additionally, if  
19 checkbox 358 is selected then interface 150 displays a "view annotations" dialog  
20 box 400 of Fig. 9. Alternatively, a user can provide a view request, causing  
21 interface 150 to display dialog box 400, by selecting show annotations button 252  
22 in annotation toolbar 240 of Fig. 6.

23 Fig. 9 shows a dialog box 400 that identifies annotations corresponding to a  
24 playlist of media segments. The playlist is a result of the query input by the user  
25 as discussed above with reference to Fig. 8. In the illustration of Fig. 9, annotation

1 identifiers in the form of user identifiers 406 and summaries 408 are displayed  
2 within an annotation listing box 402. The user can scroll through annotation  
3 identifiers in a conventional manner via scroll bars 404 and 405. The annotation  
4 identifiers are presented in annotation listing box 402 according to a default  
5 criteria, such as chronological by creation time/date, by user, alphabetical by  
6 summaries, etc.

7 Related annotations are displayed in an annotation listing 402 in a  
8 hierarchical, horizontally offset manner. The identifier of an annotation that is  
9 related to a previous annotation is "indented" from that previous annotation's  
10 identifier and a connecting line between the two identifiers is shown.

11 Dialog box 400 can be displayed concurrently with a multimedia player  
12 that is presenting multimedia content that corresponds to the annotations in  
13 annotation listing 402 (e.g., as illustrated in Fig. 10 below). Interface module 150  
14 can have the annotations "track" the corresponding multimedia content being  
15 played back, so that the user is presented with an indication (e.g., an arrow) as to  
16 which annotation(s) correspond to the current temporal position of the multimedia  
17 content. Such tracking can be enabled by selecting checkbox 422, or disabled by  
18 de-selecting checkbox 422.

19 Dialog box 400 also includes a merge annotation sets checkbox 424.  
20 Selection of merge annotation sets checkbox 424 causes interface module 150 to  
21 present annotation identifiers in listing box 402 in a chronological order regardless  
22 of what set(s) the annotations in annotation listing 402 belong to. If checkbox 424  
23 is not selected, then annotations from different sets are grouped and displayed  
24 together in annotation listing 402 (e.g., under the same tree item). Thus, when  
25

checkbox 424 is not selected, interface 150 displays one playlist for each annotation set that has been retrieved from annotation server 10.

Dialog box 400 also includes a refresh button 428, a close button 430, and an advanced button 432. Selection of refresh button 428 causes interface module 150 to communicate with annotation back end 151 to access annotation server 10 and obtain any additional annotations that correspond to the query that resulted in listing box 402.

Selection of close button 430 causes interface 150 to terminate the display of dialog box 400. Selection of advanced button 432 causes interface 150 to display a different view annotations box having additional details, such as annotation target information (analogous to target display 332 discussed below with reference to Fig. 8), user-selectable preferences for information displayed as annotation identifiers in listing box 402, etc.

Upon user selection of a particular annotation identifier from listing box 402 (e.g., "single clicking" on the summary), preview information is presented in a preview section 416, and a selection box or menu 410 is provided. The exact nature of the preview information is dependent on the data type and amount of information that was requested (e.g., as identified in level of detail 350 of Fig. 8).

Menu 410 includes the following options: play, export ASX playlist, export annotations, time order, custom order, save, and reset. Selection of the "play" option causes playback of the multimedia content to begin starting with the selected annotation in annotation list 402. Selection of the "export ASX playlist" option causes annotation backend 151 to output a record (e.g., create a file) that identifies the temporal segments of multimedia content that the annotations identified in list 402 correspond to, as determined by the begin and end times of



1 the annotations. Selection of the "export annotations" option causes annotation  
2 backend 151 to output a record (e.g., create a file) that includes the annotation  
3 content of each annotation identified in list 402.

4 Selection of the "time order" option causes interface module 150 to display  
5 the identifiers in list 402 in chronological order based on the begin time for each  
6 annotation. Selection of the "custom order" option allows the user to identify  
7 some other criteria to be used in determining the order of the identifiers in list 402  
8 (e.g., identifiers can be re-ordered in a conventional drag and drop manner). Re-  
9 ordering annotation identifiers causes the sequence numbers 204 (of Fig. 4) of the  
10 annotations to be re-ordered accordingly. Selection of the "save" option causes  
11 interface module 150 to save the current custom ordering to annotation server 10  
12 of Fig. 3 by saving the current sequence numbers of the annotations. Selection of  
13 the "reset" option causes interface module 150 to ignore any changes that have  
14 been made since the last saved custom ordering and revert to the last saved custom  
15 ordering.

16 Transfer of the corresponding media segments (and/or the annotations) to  
17 client 15 is initiated when the "play" option of menu 410 is selected. Upon  
18 selection of the play option, interface 150 of Fig. 3 provides the list of annotation  
19 identifiers being displayed to web browser 153 (or other multimedia presentation  
20 application) in the order of their display, including the target identifier and  
21 temporal range information. Thus, web browser 153 receives a list of multimedia  
22 segments that it is to present to the user in a particular order. Web browser 153  
23 then accesses media server 11 to stream the multimedia segments to client 15 for  
24 presentation in that order. By use of the play option in menu 410, a user is able to  
25 review the information regarding the annotations that satisfy his or her search



1 criteria and then modify the annotation playlist (e.g., by deleting or reordering  
2 annotation identifiers) before the corresponding media segments (and/or the  
3 annotations) are presented to him or her.

4 Alternatively, transfer of the media segments may be initiated in other  
5 manners rather than by selection of the play option in menu 410. For example, a  
6 "start" button may be included as part of dialog box 400, selection of which  
7 initiates transfer of the media segments to client 15.

8 The annotations and/or corresponding media segments are presented to the  
9 user "back to back" with very little or no noticeable gap between different  
10 annotations and between different segments. Thus, the presentation of the  
11 annotations and/or media segments is "seamless".

12 A user is able to reorder the media segments of the playlist and thereby  
13 alter their order of presentation. In the illustrated embodiment, media segments  
14 are reordered by changing the ordering of the annotation identifiers in annotation  
15 listing 402 in a drag and drop manner. For example, using a mouse and pointer a  
16 user can select a particular annotation identifier (e.g., identifier 420) and drag it to  
17 a different location within the dialog box (e.g., between identifiers 419 and 421),  
18 thereby changing when the media segment corresponding to the annotation  
19 identified by identifier 420 is presented relative to the other annotations.

20 As discussed above, information regarding the media stream as well as the  
21 particular media segment within that stream to which an annotation corresponds is  
22 maintained in each annotation. At the appropriate time, web browser 153 sends a  
23 message to the appropriate media server 11 of Fig. 1 to begin streaming the  
24 appropriate segment to client computer 15. Web browser 153, knowing the  
25 duration of each of the segments being provided to client computer 15, forwards

1 additional messages to media server 11 to continue with the provision of the next  
2 segment, according to the playlist, when appropriate. By managing the delivery of  
3 the media segments to client computer 15 in such a manner, web browser 153 can  
4 keep the media segments being provided to the user in a seamless manner.

5 According to an alternate embodiment, the media segments could be  
6 streamed to annotation server 10 for temporary buffering and subsequent  
7 streaming to client computer 15. According to another alternate embodiment,  
8 identifying information (e.g., source, start time, and end time) for the media  
9 segment could be provided to media server 11 from annotation server 10 for  
10 streaming to client computer 15.

11 Additionally, according to one embodiment the collection of media  
12 segments identified by the playlist can be stored as an additional media stream by  
13 selecting "export ASF playlist" option in menu 410 of Fig. 9. By saving the  
14 collection of media segments as a single media stream, the collection can be  
15 retrieved by the user (or other users) at a later time without having to go through  
16 another querying process. Furthermore, the collection of segments, stored as a  
17 media stream, can itself be annotated.

18 The collection of segments can be stored as a media stream in any of a  
19 variety of different locations and formats. The media stream can be stored in an  
20 additional data store (not shown) managed by annotation server 10 of Fig. 3, or  
21 alternatively stored at media server 11 of Fig. 1 or another media server (not  
22 shown) of Fig. 1. According to one embodiment, the media stream includes the  
23 source information, start time, and end time for each of the segments in the  
24 playlist. Thus, little storage space is required and the identifying information for  
25 each of the segments is independent of the annotations. Alternatively, the media

1 stream includes pointers to each of the annotations. For subsequent retrieval of the  
2 media segments, the stored pointers can be used to retrieve each of the appropriate  
3 annotations, from which the corresponding media segments can be retrieved.  
4 According to another alternate embodiment, the media segments themselves could  
5 be copied from media server 11 of Fig. 1 and those segments stored as the media  
6 stream.

7 Fig. 10 shows one implementation of a graphical user interface window 450  
8 that concurrently displays annotations and corresponding media segments. This  
9 UI window 450 has an annotation screen 454, a media screen 456, and a toolbar  
10 240.

11 Media screen 456 is the region of the UI within which the multimedia  
12 content is rendered. For video content, the video is displayed on screen 456. For  
13 non-visual content, screen 456 displays static or dynamic images representing the  
14 content. For audio content, for example, a dynamically changing frequency wave  
15 that represents an audio signal is displayed in media screen 456.

16 Annotation screen 454 is the region of the UI within which the annotation  
17 identifiers and/or annotation content are rendered. For example, dialog box 400 of  
18 Fig. 9 can be annotation screen 454.

19 Fig. 11 illustrates methodological aspects of one embodiment of the  
20 invention in retrieving and presenting annotations and media segments to a user.

21 A step 500 comprises displaying a query dialog box 330 of Fig. 8.  
22 Interface 150 of Fig. 3 provides dialog box 330 in response to a query request  
23 from a user, allowing the user to search for annotations that satisfy various user-  
24 definable criteria.

1 A step 502 comprises receiving query input from the user. Interface 150 of  
2 Fig. 3 receives the user's input(s) to the query dialog box and provides the inputs  
3 to annotation server 10 of Fig. 3.

4 A step 504 comprises generating an annotation list. ABE 132 of Fig. 3 uses  
5 the user inputs to the query dialog box to select annotations from stores 17 and 18.  
6 ABE 132 searches through annotation meta data store 18 for the annotations that  
7 satisfy the criteria provided by the user. The annotations that satisfy that criteria  
8 then become part of the annotation list and identifying information, such as the  
9 annotation titles or summaries, are provided to client 15 by annotation server 10.

10 A step 506 comprises displaying a view annotations dialog box 400 of Fig.  
11 9 that contains the annotation identifying information from the annotation list  
12 generated in step 504.

13 Steps 508 and 510 comprise receiving user input selecting various  
14 annotations from the identifying information displayed in step 506. Steps 508 and  
15 510 repeat until the user has finished his or her selecting.

16 A step 512 comprises retrieving the selected annotations and corresponding  
17 media segments. ABE 132 in annotation server 10 of Fig. 3 is responsible for  
18 retrieving the selected annotations from stores 17 and 18.

19 A step 514 comprises presenting the selected annotations and  
20 corresponding media segments to the user in a seamless manner.

21 In the illustrated embodiment, both the selected annotations as well as the  
22 corresponding media segments are provided to the user. In one alternate  
23 embodiment, only the media segments corresponding to the annotations (and not  
24 the annotations themselves) are provided to the user. In another alternate  
25 embodiment only the annotations (and not the corresponding segments of the

1 media stream) are provided to the user. In another embodiment, the annotations  
2 are downloaded to the client computer first, and the media segments are  
3 downloaded to the client computer later in an on-demand manner.

4 In the illustrated embodiment, annotation data is buffered in annotation  
5 server 10 of Fig. 1 for provision to client 15 and media stream data is buffered in  
6 media server 11 for provision to client 15. Sufficient buffering is provided to  
7 allow the annotation and media stream data to be provided to the client seamlessly.  
8 For example, when streaming two media segments to client 15, as the end of the  
9 first media segment draws near media server 11 is working on obtaining and  
10 streaming the beginning of the second media segment to client 15. By doing so,  
11 there is little or no noticeable gap between the first and second media segments as  
12 presented to the user. Alternatively, rather than providing such buffering in the  
13 servers 10 and 11, additional buffering can be provided by client 15 to allow the  
14 seamless presentation of the data.

## 15 Conclusion

16 The invention described above provides interactive playlist generation  
17 using annotations. Annotations correspond to media segments of one or more  
18 multimedia streams. Identifiers of the annotations can be presented to a user and  
19 their order advantageously changed by the user to suit his or her desires. These  
20 annotation identifiers correspond to a playlist of media segments. The media  
21 segments (and/or the annotations themselves) of the playlist are then  
22 advantageously presented to the user in a seamless, contiguous manner.  
23

24 Although the invention has been described in language specific to structural  
25 features and/or methodological steps, it is to be understood that the invention

1 defined in the appended claims is not necessarily limited to the specific features or  
2 steps described. Rather, the specific features and steps are disclosed as preferred  
3 forms of implementing the claimed invention.  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25

1 **CLAIMS**

2 1. One or more computer-readable media containing a computer  
3 program for annotating streaming media, wherein the program performs steps  
4 comprising:

5 creating annotations interactively with a user, wherein the annotations  
6 correspond to identified segments of one or more media streams;

7 graphically ordering the annotations in a desired order of presentation in  
8 response to user input; and

9 in response to a user instruction, sequentially presenting the annotations  
10 along with their corresponding identified media stream segments in the desired  
11 order of presentation.

12  
13 2. One or more computer-readable media as recited in claim 1, wherein  
14 the annotations comprise textual annotations.

15  
16 3. One or more computer-readable media as recited in claim 1, wherein  
17 the media streams comprise audio/visual video streams.

18  
19 4. One or more computer-readable media as recited in claim 1, wherein:  
20 the annotations are textual annotations;  
21 the media streams are audio/visual video streams; and  
22 the presenting step comprises displaying the textual annotations in one  
23 display area while displaying the corresponding segments of the audio/visual  
24 streams in another display area.  
25

1           5.     One or more computer-readable media as recited in claim 1, the steps  
2 further comprising storing the annotations and their desired order of presentation.

3  
4           6.     One or more computer-readable media as recited in claim 1, the steps  
5 further comprising:

6           storing the annotations and their desired order of presentation; and

7           in response to a user request,

8                 retrieving the stored annotations and their desired order of  
9 presentation,

10                displaying the retrieved annotations in their desired order of  
11 presentation, and

12                retrieving and presenting the media stream segments identified by  
13 the retrieved annotations, in sequential order in accordance with the desired  
14 order of presentation of the retrieved annotations.

15  
16           7.     A method comprising:

17           receiving an indication of a plurality of annotations selected by a user,  
18 wherein each of the plurality of annotations corresponds to a media stream or to  
19 one or more media streams; and

20           seamlessly providing one or more of,

21                 the plurality of annotations, and

22                 at least a portion of the media stream corresponding to each of the  
23 plurality of annotations.



1           8.     A method as recited in claim 7, wherein the seamlessly providing  
2 comprises providing the plurality of annotations and the portions of the media  
3 streams corresponding to the plurality of annotations to a client computer for  
4 seamless presentation to a user.

5  
6           9.     A method as recited in claim 7, wherein each of the plurality of  
7 annotations corresponds to a segment of one of the one or more media streams,  
8 each segment being less than the entire stream.

9  
10          10.    A method as recited in claim 7, wherein the seamlessly providing  
11 comprises:

12               seamlessly providing the plurality of annotations concurrently with  
13 seamlessly providing at least a portion of the media stream corresponding to each  
14 of the plurality of annotations.

15  
16          11.    A method as recited in claim 7, further comprising:  
17               presenting a plurality of annotation identifiers to the user; and  
18               wherein the seamlessly providing comprises seamlessly providing the one  
19 or more of the plurality of annotations and the portion of the media stream  
20 corresponding to each of the plurality of annotations in an order defined by the  
21 order of the plurality of annotation identifiers.

22  
23          12.    A method as recited in claim 11, further comprising:  
24               allowing the ordering of the plurality of annotation identifiers to be  
25 changed by the user.

1  
2       **13.**    A method as recited in claim 12, further comprising:  
3       allowing the user to change the order of the plurality of annotation  
4 identifiers in a drag and drop manner.

5  
6       **14.**    A method as recited in claim 7, further comprising:  
7       storing the at least a portion of the media stream corresponding to each of  
8 the plurality of annotations as a new media stream of the one or more media  
9 streams.

10  
11       **15.**    A method as recited in claim 7, wherein each of the plurality of  
12 annotations comprises one or more of audio data and text data.

13  
14       **16.**    A method as recited in claim 7, wherein each of the one or more  
15 media streams comprises audio and video data.

16  
17       **17.**    A computer-readable memory containing a computer program that is  
18 executable by a computer to perform the method recited in claim 7.

19  
20       **18.**    A system comprising:  
21       an annotation database that stores one or more collections of annotations,  
22 wherein each of the annotations identifies at least a segment of a media stream;  
23 and  
24  
25

1 an annotation module to control storage and retrieval of the plurality of  
2 annotations, wherein the annotation module is configured to perform steps  
3 comprising:

4 retrieving a particular collection of annotations from the annotation  
5 database;

6 presenting the annotations of the retrieved collection to a user; and  
7 managing sequential presentation to the user of the media stream  
8 segments corresponding to the presented annotations.

9  
10 **19.** A system as recited in claim 18, wherein the annotation module is  
11 further configured to perform a step of communicating with a client computer to  
12 provide indications of the plurality of annotations to the client computer for  
13 display to the user.

14  
15 **20.** A system as recited in claim 19, wherein the indications of the  
16 plurality of annotations comprise summary information for each of the plurality of  
17 annotations.

18  
19 **21.** A system as recited in claim 19, wherein each of the plurality of  
20 annotations corresponds to an annotation set, and wherein the annotation module  
21 is further configured to perform a step of providing the annotation set information  
22 to the client computer.

1           **22.**    A system as recited in claim 18, wherein each of the media stream  
2 segments comprises audio and video data.

3  
4           **23.**    A system as recited in claim 18, wherein the annotation module is  
5 further configured to perform a step of saving information regarding the media  
6 stream segments as an additional new media stream.

7  
8           **24.**    A system as recited in claim 23, wherein the information regarding  
9 each of the media stream segments comprises an identifier of a media stream of  
10 which the media segment is a part, a temporal location in the media stream  
11 identifying where the media segment begins, and a temporal location in the media  
12 stream identifying where the media segment ends.

13  
14           **25.**    A system as recited in claim 18, further comprising:  
15           a client computer, coupled to the annotation module, configured to receive  
16 the media stream segments and present the media stream segments to the user.

17  
18           **26.**    A system as recited in claim 25, further comprising:  
19           a media server, coupled to the annotation module, having access to a  
20 plurality of media streams, the media server configured to provide at least a  
21 portion of the plurality of media streams to the client computer as the media  
22 stream segments.

1       **27.**     A system as recited in claim 18, wherein each of the plurality of  
2 annotation identifiers corresponds to a single media stream of the plurality of  
3 media streams.

4  
5       **28.**     One or more computer-readable storage media containing a program  
6 having instructions that are executable by a computer to perform steps comprising:  
7       configuring a first portion of a user interface to display a plurality of  
8 identifiers corresponding to a plurality of annotations, the plurality of identifiers  
9 corresponding to a playlist of media segments to be seamlessly presented to a user;  
10 and

11       reordering the plurality of identifiers in accordance with user input to  
12 change the order in which the media segments are to be presented.

13  
14       **29.**     One or more computer-readable storage media as claimed in claim  
15 28, the program having instructions that are executable by the computer to further  
16 perform a step comprising:

17       receiving the media segments from a media server in an order determined  
18 by the playlist.

19  
20       **30.**     One or more computer-readable storage media as claimed in claim  
21 28, the program having instructions that are executable by the computer to further  
22 perform steps comprising:

23       receiving the media segments from a media server in an order determined  
24 by the playlist; and  
25

1 presenting the media segments at the user interface in the order determined  
2 by the playlist.  
3

4 **31.** One or more computer-readable storage media as claimed in claim  
5 28, the program having instructions that are executable by the computer to further  
6 perform a step comprising:

7 allowing the user to reorder the plurality of identifiers in a drag and drop  
8 manner.  
9

10 **32.** One or more computer-readable storage media as claimed in claim  
11 28, the program having instructions that are executable by the computer to further  
12 perform a step comprising:

13 configuring a second portion of the user interface to present the plurality of  
14 annotations concurrently with the media segments.  
15

16 **33.** One or more computer-readable storage media as claimed in claim  
17 28, wherein each of the media segments comprises audio and video data.  
18  
19  
20  
21  
22  
23  
24  
25

## ABSTRACT

A plurality of user-selected annotations are used to define a playlist of media segments corresponding to the annotations. The user-selected annotations and their corresponding media segments are then provided to the user in a seamless manner. A user interface allows the user to alter the playlist and the order of annotations in the playlist. The user interface identifies each annotation by a short subject line.

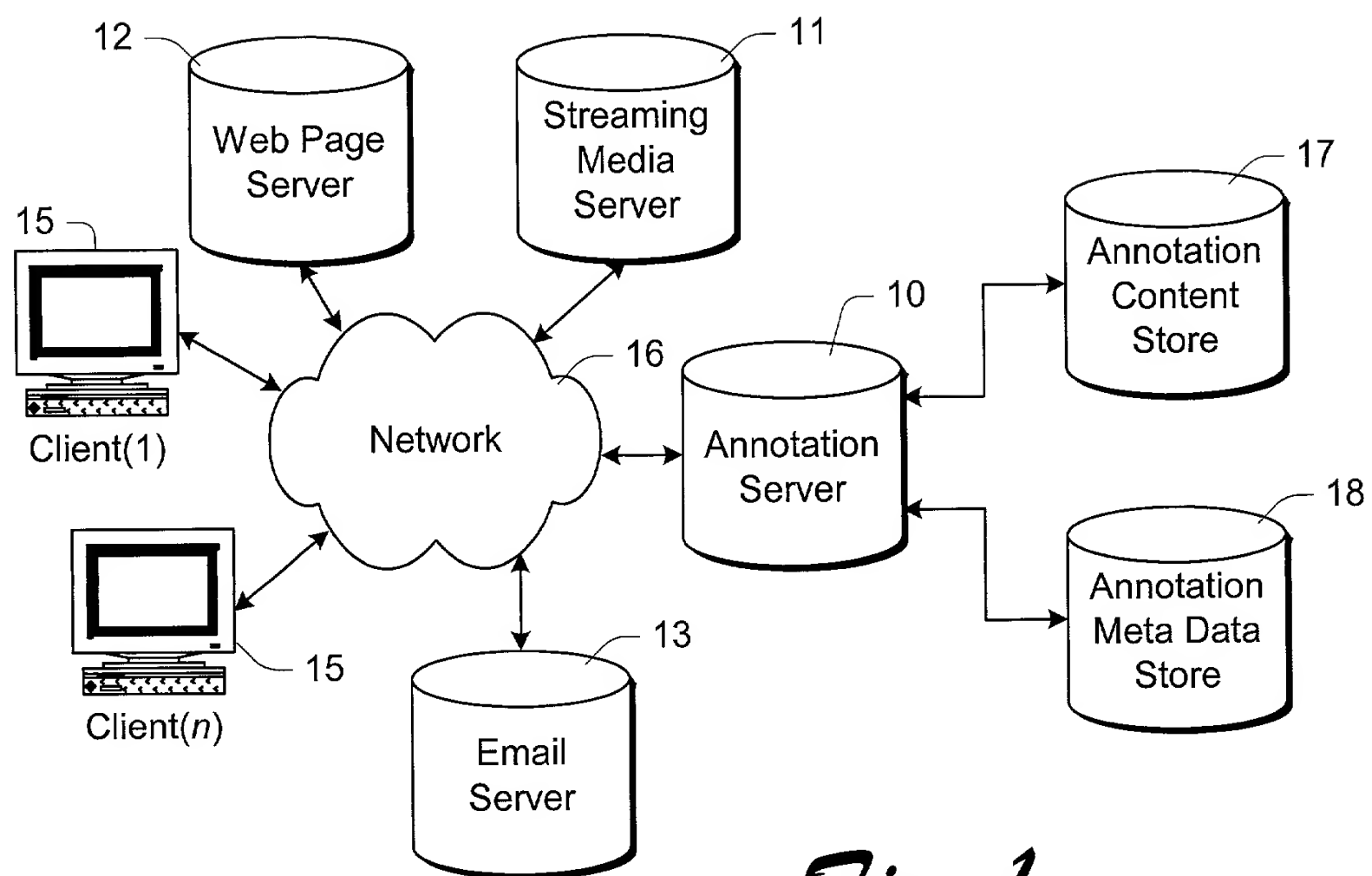


Fig. 1

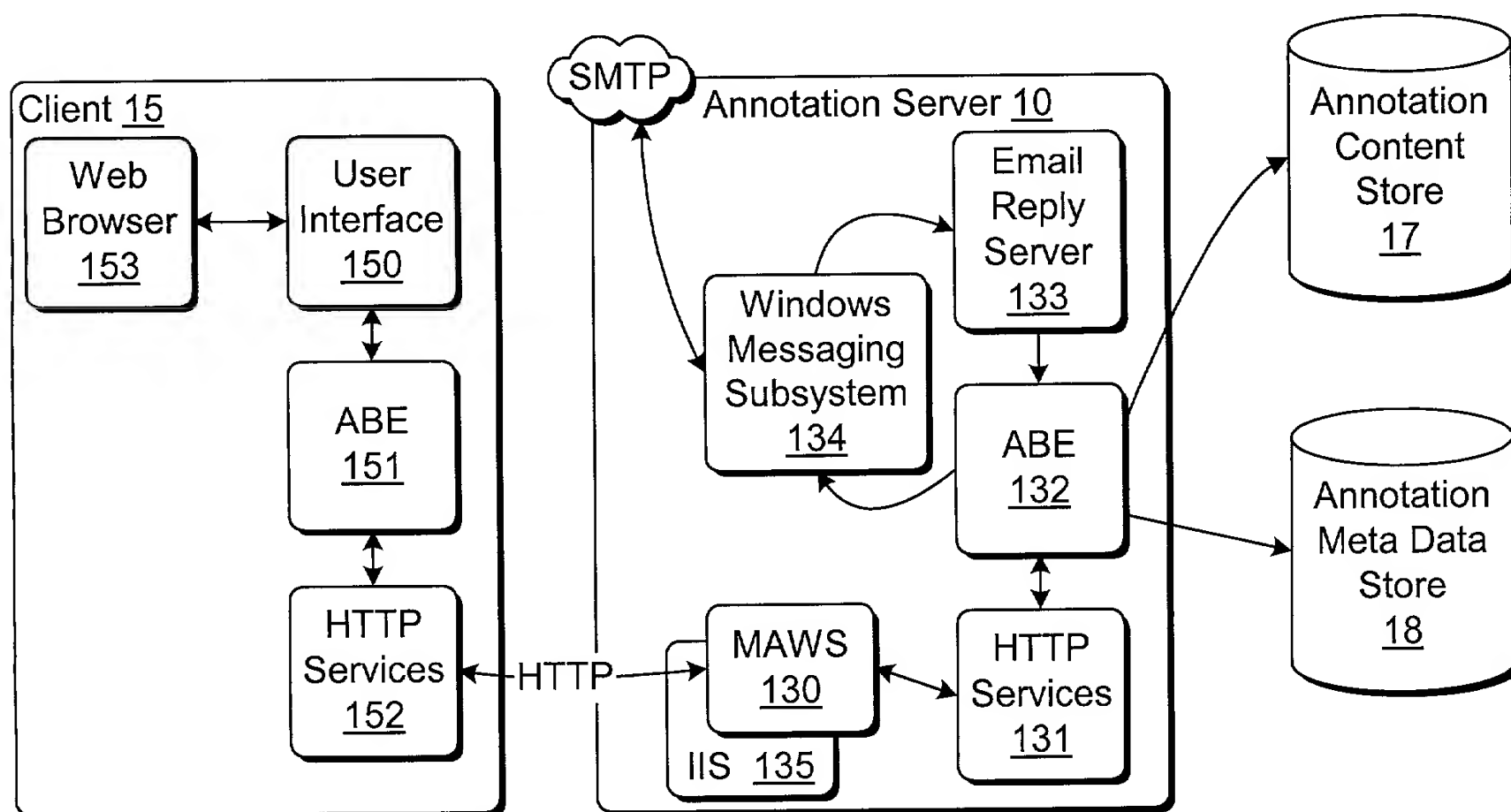
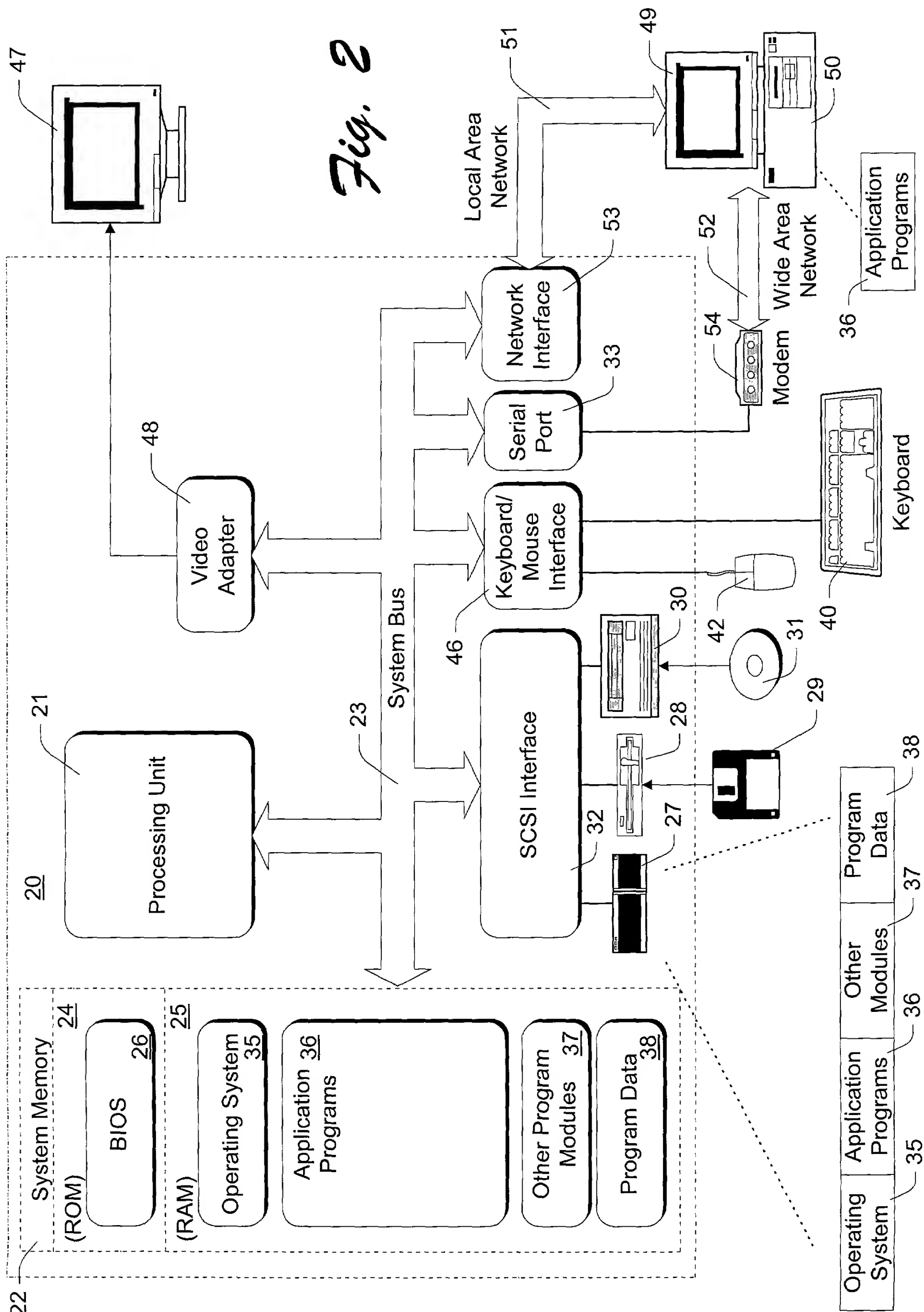


Fig. 3



Fig. 2



Annotation Entry  
180 ↘

Author <u>182</u>
Time Range <u>184</u>
Time Units <u>186</u>
Creation Time <u>188</u>
Title <u>190</u>
Content <u>192</u>
Annotation Identifier <u>194</u>
Related Annotation Identifier <u>196</u>
Set Identifier(s) <u>198</u>
Media Content Identifier <u>200</u>
Property Fields <u>202</u>
Sequence Number <u>204</u>

Fig. 4

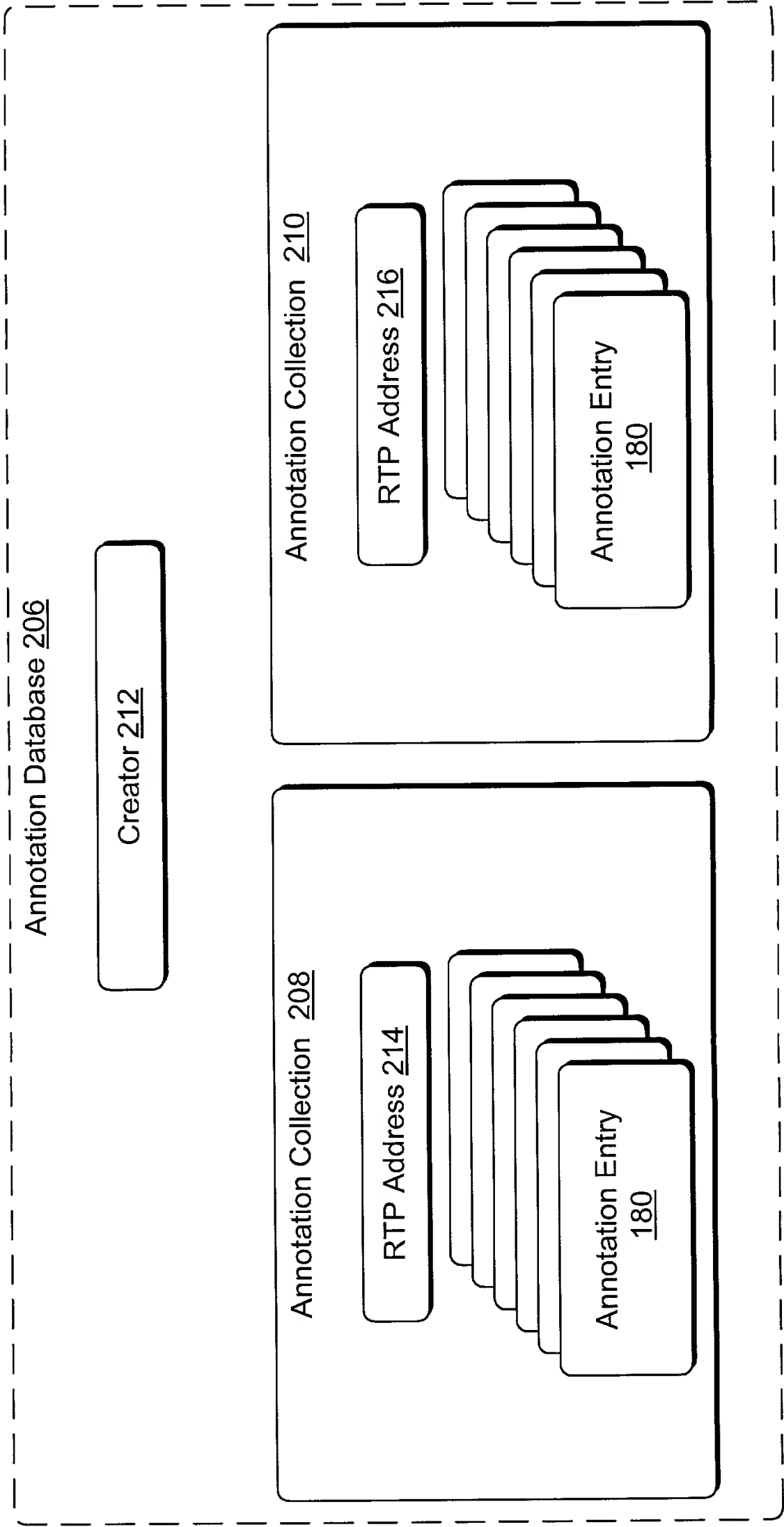


Fig. 5

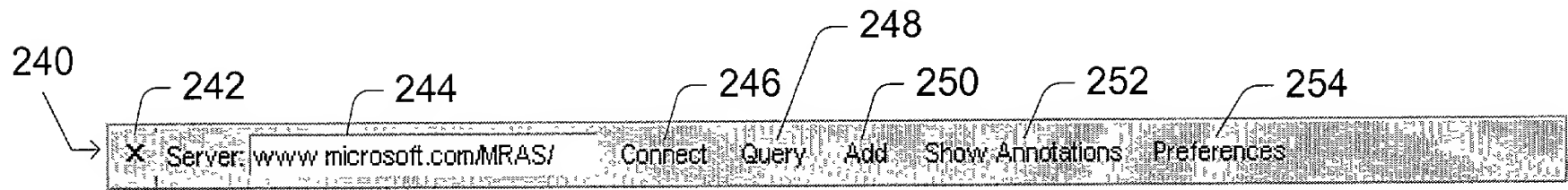


Fig. 6

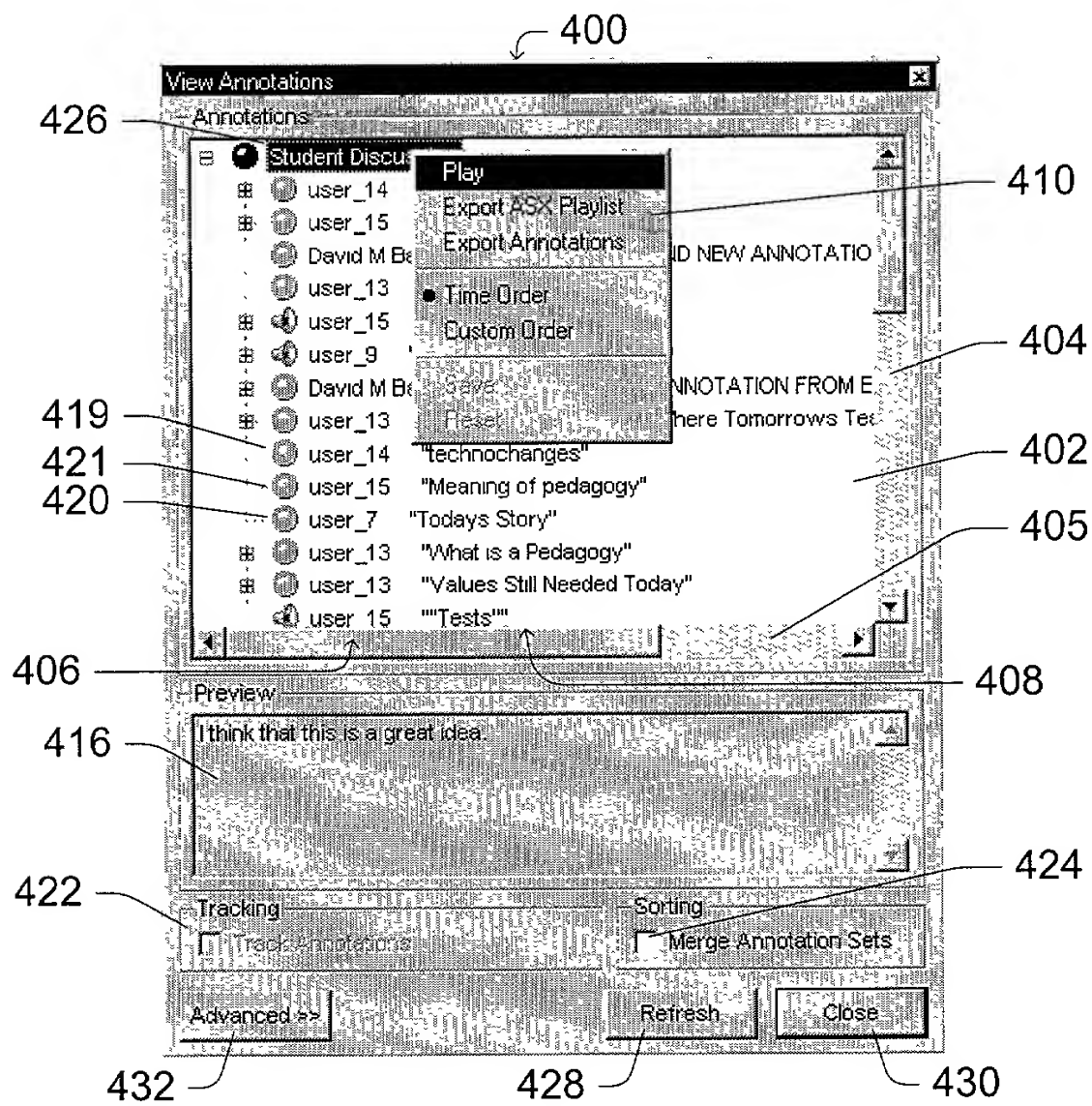


Fig. 9

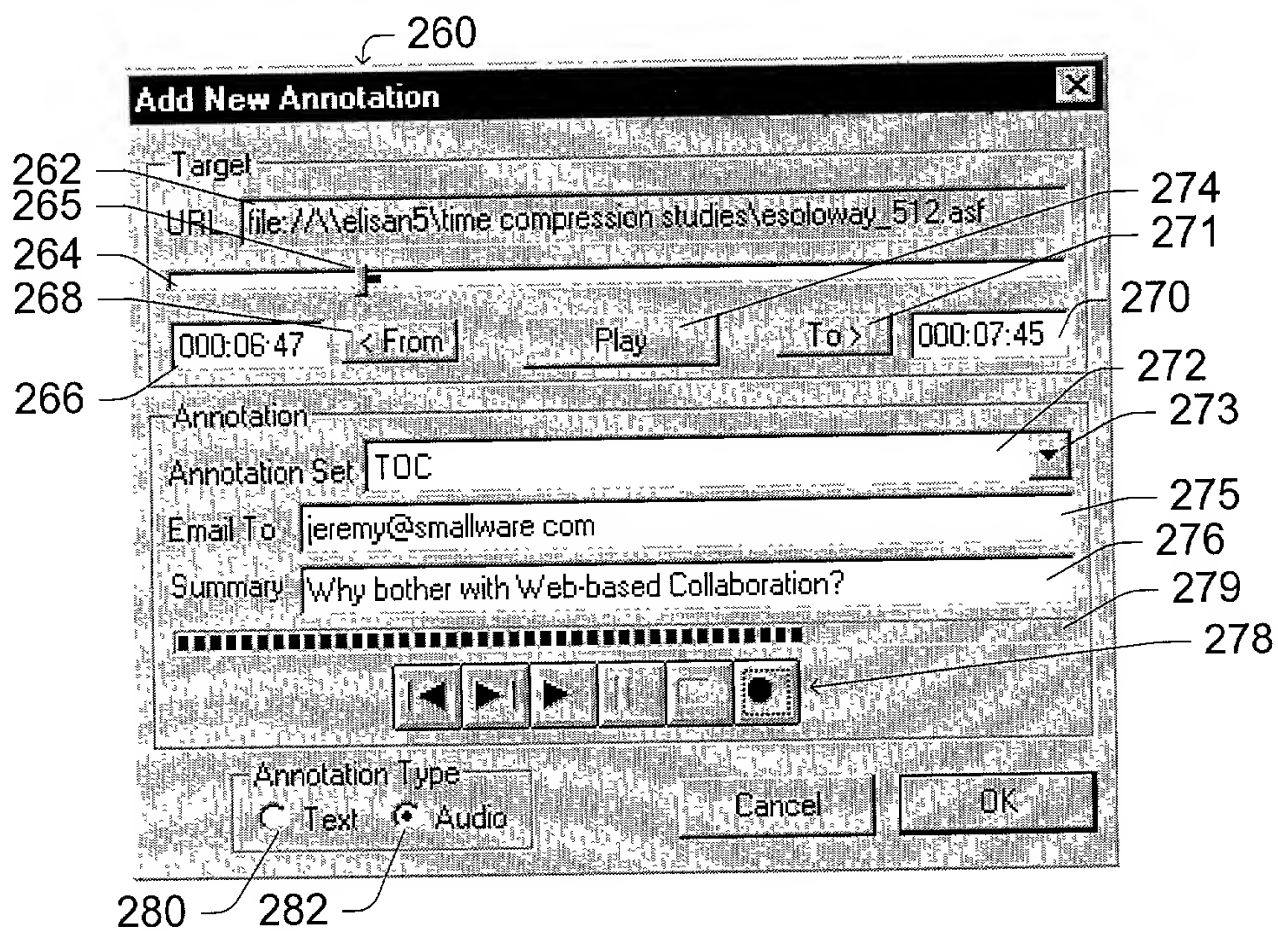


Fig. 7

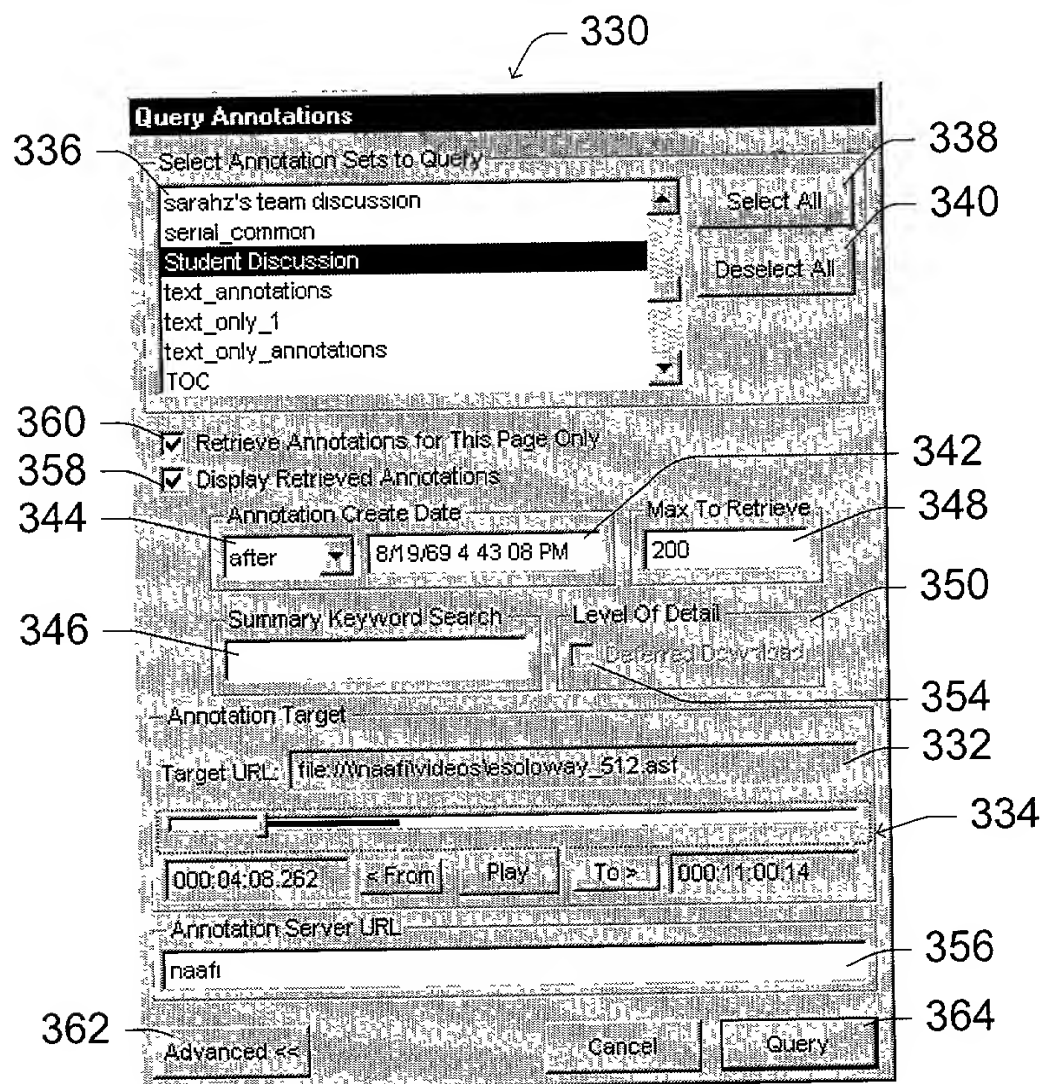


Fig. 8

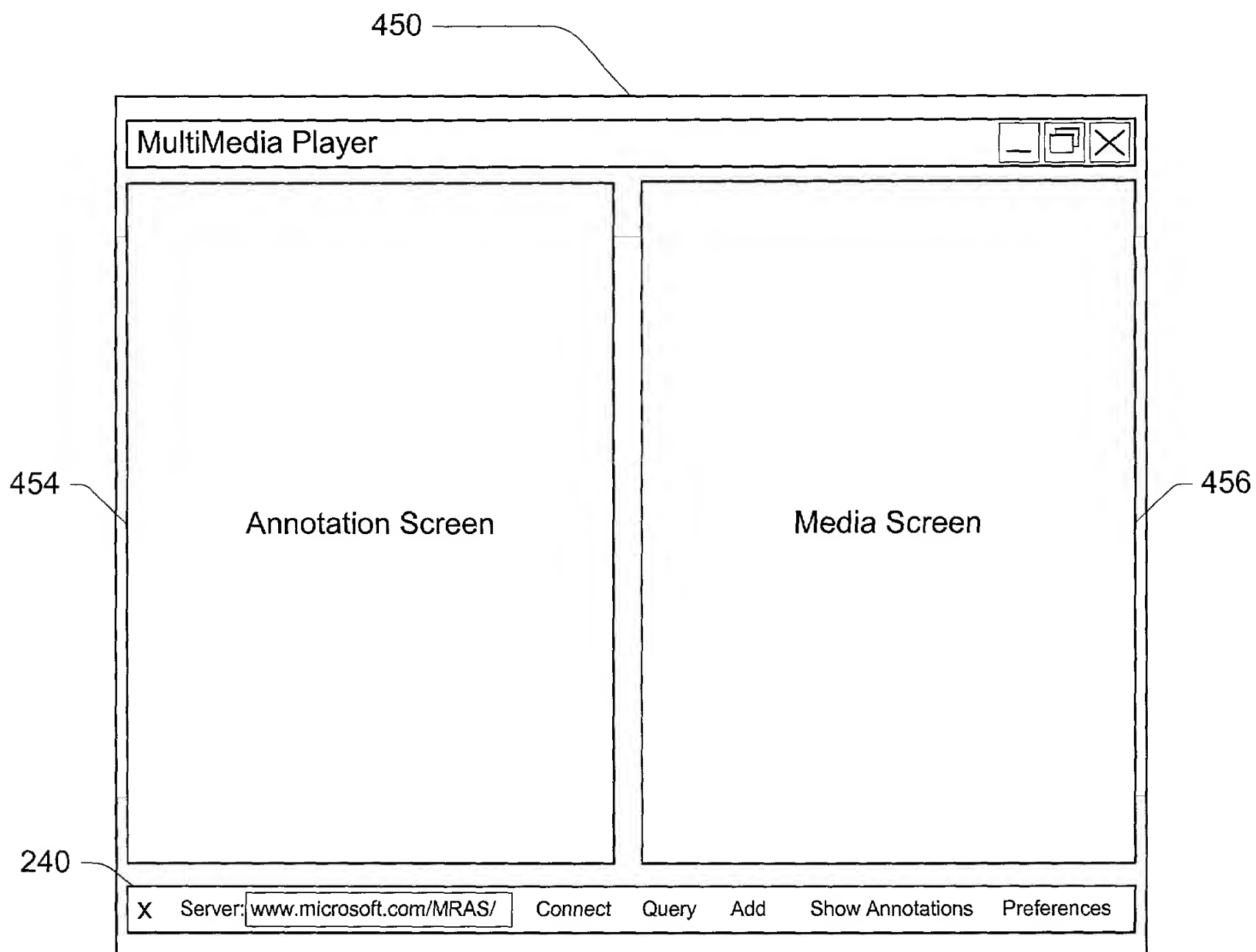
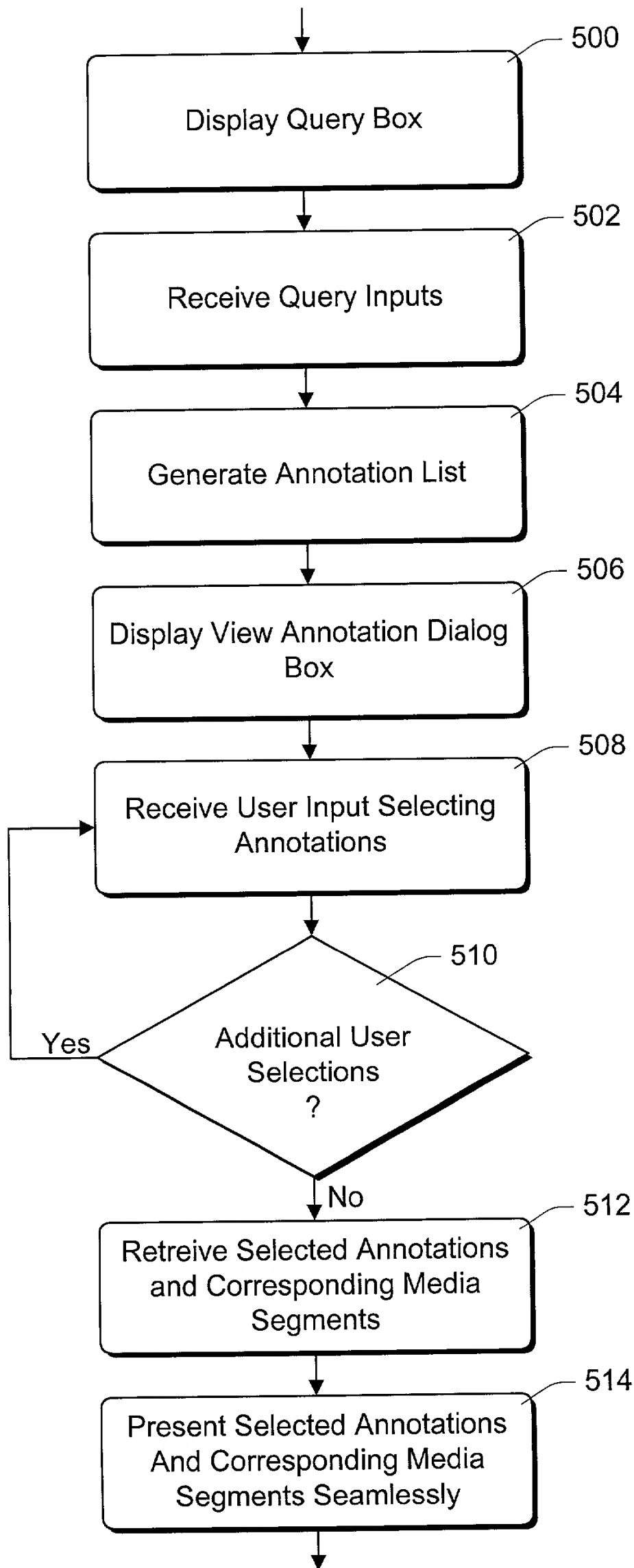


Fig. 10

*Fig. 11*